Joint Semi-blind Detection and Channel Estimation in Space-Frequency Trellis Coded MIMO-OFDM

Robert J. Piechocki, Andrew R. Nix and Joe P. McGeehan

University of Bristol, Centre For Communications Research
Woodland Road, MVB, BS8 1UB, Bristol, UK.
E-mail: r.j.piechocki@bristol.ac.uk, Tel. +44 117 9545203, Fax. +44 117 9545206

***Abstract*— This paper considers an OFDM system with a Multiple-Input Multiple-Output (MIMO) configuration, which uses Space-Frequency Trellis Coding (SFTC). A novel method of decoding SFTC without a need to transmit separate training sequences is developed. The technique uses only a single frequency tone to acquire a complete set of the channels' estimates while performing SFTC detection. The method is akin to blind trellis search techniques (per-survivor processing - PSP) and adaptive Viterbi. Our solution consists of the deployment of a bank of Kalman Filters. The bank of Kalman Filters is coupled with Viterbi type decoders, which produce tentative decisions based on Kalman channel predictions. In return, the Kalman filters use the tentative decisions to update and track the MIMO channels corresponding to a number of tracked hypotheses.**

***Index Terms*— Space-Frequency Coding, MIMO-OFDM, Semi-blind detection, Kalman Filter, Per-Survivor Processing, Adaptive Viterbi**

I. INTRODUCTION

THE current generation of high data rate wireless local area network (WLAN) standards, such as IEEE802.11a, provide data rates of up to 54 Mbit/s. However, the ever-increasing demand for even higher data rate services, such as internet, video and multi-media, have created a need for improved bandwidth efficiency from next generation wireless LANs. The current IEEE802.11a standard employs the bandwidth efficient scheme of Orthogonal Frequency Division Multiplex (OFDM) and adaptive modulation and demodulation. The system was designed as single-input single-output (SISO) systems, essentially employing a single transmit and receive antenna at each end of the link. Within ETSI BRAN some provision for multiple antennas, or sectorised antennas, has been investigated for improved diversity gain and thus link robustness.

Until recently considerable effort was put into designing systems so as to mitigate the perceived detrimental effects of multipath propagation, which are especially prevalent in indoor wireless LAN environments. However, recent work [1] has shown that by utilising multiple antenna architectures at both the transmitter and receiver, so-called muliple-input multiple-output (MIMO) architectures, vastly increased channel capacities are possible. The ideas behind space-time trellis coded modulation (STTCM) were first presented in [2]. Recent attention has turned to the adoption of space-time coding techniques to wideband channels, and in particular their usage in OFDM-based systems where coding is performed in the space-frequency domain [3]. The maximum likelihood detection of SFTC requires provision of the channel state information (CSI). Typically the CSI is acquired via training sequences (e.g. both Hiperlan/2 and IEEE 802.11a standards include transmission of preambles for this purpose). The resulting CSI estimates are then fed to a space-frequency Viterbi decoder, which performs a MLSE search.

Kalman filter tracking of a Space-time block coded system was first reported in [4]. An attempt has also been made to jointly estimate and decode space-time trellis codes [5]. In [5] the authors propagate *posterior* distribution for the decoded symbols. This results in a complex nonlinear problem that can only be tackled by particle filters (sequential importance sampling). However, this also introduces a phase ambiguity in this application. Our approach is computationally less demanding and avoids the phase ambiguity problems even in a blind version. The approach is based on per-survivor processing (PSP) (blind trellis search techniques) proposed in [6] and [7], to cope with the problem of unknown or fast changing channels. The PSP techniques are based on adaptive Viterbi and LMS detection, and were developed in the context of blind MLSE equalisation. However, PSP is unsuitable in MIMO case.

The novel technique uses only a few frequency tones to acquire partial estimate of the channels - hence the term semi-blind. In fact, it is shown that only one pilot tone (training on one harmonic) provides sufficient information to reliably decode and estimate the entire frame and all CSI. Such pilot tones are already incorporated in most OFDM systems to correct for residual phase estimation errors (in addition to training sequences). The proposed technique naturally lends itself to a blind scheme, where even partial knowledge of the channels is not essential. Our solution relies on a bank of Kalman Filters. The bank of Kalman Filters is coupled with Viterbi type decoders, which produce tentative decisions based on Kalman channel predictions. In return the bank of Kalman filters uses the tentative decisions to update and track the MIMO channels corresponding to the tracked hypothesis.

This paper is structured as follows: Section II discusses space-frequency coding for MIMO-OFDM and sets up the signal model. In section III the proposed technique is developed, preceded by a review of recursive estimation. Section IV presents simulation results and section V provides conclusions.

II. Space-frequency coded OFDM

When STTCM is applied to OFDM systems the coding takes place across frequency and space rather than time and space - figure 1. In the time domain the amount of available diversity is related to the Doppler spread of a signal. Hence for low mobility high data rate systems (as considered here), the channel remains almost constant over a frame. However, the delay spread in the radio channel gives rise to diversity in the frequency domain.

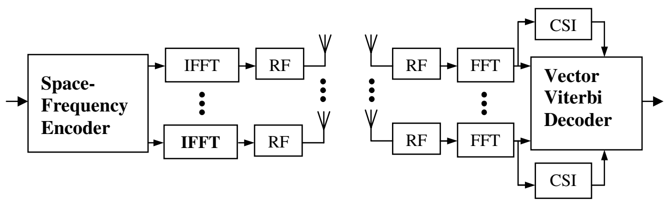

Fig. 1. Space-frequency coded MIMO-OFDM system.

An OFDM system with a Cyclic Prefix (CP-OFDM) is assumed here, where the i^{th} transmitted block of data $\bar{\mathbf{u}}_i$ is given by $\bar{\mathbf{u}}_i = \mathbf{T}_{CP}\mathbf{F}^{-1}\mathbf{u}_i$. The data vector, $\mathbf{u}_i$, is of length K, the size of the CP insertion matrix, $\mathbf{T}_{CP}$, is $P \times K$, where $P = C + K$, C represents the length of the CP, and the Fourier transform matrix $\mathbf{F}$ is of size $K \times K$. The receiver receives the current transmitted block of data $\bar{\mathbf{u}}_i$, in addition to a fraction of the previous block through the excess length of the channel impulse response. This is described by Toeplitz channel matrices $\mathbf{H}_0$ and $\mathbf{H}_1$, and the received signal block $\bar{\mathbf{x}}_i$ pertaining to $\mathbf{u}_i$ is given by:

$$\bar{\mathbf{x}}_i = \mathbf{H}_0\bar{\mathbf{u}}_i + \mathbf{H}_1\bar{\mathbf{u}}_{i-1} + \bar{\eta}_i \tag{1}$$

Both of the above channel matrices are of size $P \times P$ and are given by: $(h_0 \ \cdots \ h_{L-1} \ 0 \ \cdots 0\ )^T$ for the first column and $(h_0 \ 0 \ \cdots \ 0\ )$ for the first row of $\mathbf{H}_0$; $(0 \ \cdots \ 0\ )^T$ for the first column and $(0 \ \cdots \ h_{L-1} \ \cdots \ h_1)$ for the first row of $\mathbf{H}_1$. Additionally, it is assumed that the length of the CP is $C \geq L-1$. The receiver removes the first C entries of $\bar{\mathbf{y}}_i$ that are affected by IBI (Inter-Block-Interference). This is achieved by pre-multiplication with a matrix $\mathbf{T}_R$ defined as: $\mathbf{T}_R = [\mathbf{0}_{K\times C}, \mathbf{I}_{K\times K}]$. Hence the input-output relationship can be expressed as:

$$\mathbf{x}_i = \mathbf{F}\mathbf{T}_R\mathbf{H}_0\mathbf{T}_{CP}\mathbf{F}^{-1}\mathbf{u}_i + \mathbf{F}\eta_i \tag{2}$$

where η_i represents the ubiquitous additive noise vector.

A specific construction of $\mathbf{T}_{CP}$ guarantees that the concatenation $\mathbf{T}_R\mathbf{H}_0\mathbf{T}_{CP}$ is a circulant matrix, and thus is diagonalised by $\mathbf{F}$. Hence: $\mathbf{F}\mathbf{T}_R\mathbf{H}_0\mathbf{T}_{cp}\mathbf{F}^{-1} = \mathbf{\Lambda} = diag\left\{\lambda^{(1)}, \cdots, \lambda^{(K)}\right\}$ and now:

$$\mathbf{x} = \mathbf{\Lambda}\mathbf{u} + \mathbf{F}\eta \tag{3}$$

The block index i has been dropped, since IBI has been alleviated.

We adopt the generating matrix representation of the STTCM codes and the encoding process [8]. This is now extended to the case of STC-OFDM. The frequency-space codeword $\mathbf{c}$ at the subcarrier k is obtained from:

$$\mathbf{c}_k = \mathcal{M}\left(\mathbf{d}_k\mathbf{G}\left(\mathrm{mod} M\right)\right) \tag{4}$$

Where: $\mathbf{d}_k = \left(\begin{array}{ccccc} d_{mk+(m-1)} & \cdots & d_{mk} & \cdots & d_{mk-s}\end{array}\right)$ denotes the $m+s$ long stream of input bits. In order to facilitate the description of an FSC-OFDM system with N_T transmit and N_R received antennas, a set of K $\{\mathbf{H}_{1:K}; k \in N\}$ $N_R \times N_T$ dimensional matrices are defined:

$$\mathbf{H}_k = \left[\begin{array}{cccc} \lambda_{1,1}^{(k)} & \lambda_{1,2}^{(k)} & \cdots & \lambda_{1,N_T}^{(k)} \\ \lambda_{2,1}^{(k)} & \lambda_{2,2}^{(k)} & \cdots & \lambda_{2,N_T}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{N_R,1}^{(k)} & \lambda_{N_R,2}^{(k)} & \cdots & \lambda_{N_R,N_T}^{(k)} \end{array}\right] \tag{5}$$

Where $\lambda_{m,n}^{(k)}$ represents the frequency response of a channel between the n^{th} transmit and the m^{th} receive antenna at the k^{th} subcarrier. Similarly, denoting $\mathbf{y}_k = \left[x_1^{(k)}, \ldots, x_{N_R}^{(k)}\right]^T$ it can be shown that the received signal at the k^{th} sub-carrier has a form:

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{c}_k + \mathbf{n}_k \tag{6}$$

In an AWGN channel, the maximum likelihood decoder can be realised using the Viterbi algorithm with the Euclidean metric:

$$\tilde{\mathbf{C}} = \underset{\tilde{\mathbf{C}}=\tilde{\mathbf{c}}_1\ldots\tilde{\mathbf{c}}_K}{\arg\min} \sum_{k=1}^{K} \left\| \mathbf{y}_k - \hat{\mathbf{H}}_k\tilde{\mathbf{c}}_k \right\|^2 \tag{7}$$

In Equation (7) it is assumed that some form of channel estimation resulting in $\left\{\hat{\mathbf{H}}_{1:K}\right\}$ has been performed prior to detection.

III. Recursive channel estimation in Space-frequency trellis coded systems

Is this section the novel technique is developed. As aforementioned we will assume that only $\left\{\hat{\mathbf{H}}_0\right\}$ is available and that this knowledge suffices to estimate $\left\{\hat{\mathbf{H}}_{1:K}\right\}$ and decode the SFTC code. Since the concept rests on the developments from recursive Bayesian estimation theory, we start with a brief recap on recursive Bayesian estimation.

A. Recursive Bayesian estimation

In Bayesian estimation some statistical knowledge about the estimated data or parameters is assumed to be available before the actual measurements take place. This knowledge is expressed in a form of a joint *a priori* probability density function. Even a decision can be made before the measurements - perhaps as a mean or a mode of the *a priori* density. In recursive estimation it is assumed that the estimated problem evolves (typically in time, however here it is in the frequency domain) and it is logical to make the decisions sequentially, which also reduces the computational effort.

The hidden states (unobserved signal) of interest $\{\mathbf{h}_k; k \in N\}$ are modelled as a Markov process:

$$f(\mathbf{h}_k | \mathbf{h}_0, \ldots, \mathbf{h}_{k-1}, \mathbf{y}_1, \ldots, \mathbf{y}_k) = f(\mathbf{h}_k | \mathbf{h}_{k-1}) \quad (8)$$

The observations $\{\mathbf{y}_{1:k}\}$ are independently and identically distributed (iid) conditioned on the current state $f(\mathbf{y}_k | \mathbf{h}_0, \ldots, \mathbf{h}_k, \mathbf{y}_1, \ldots, \mathbf{y}_{k-1}) = f(\mathbf{y}_k | \mathbf{h}_k)$.

At frequency index k the joint posterior distribution is given by Bayes' theorem:

$$f(\mathbf{h}_{0:k} | \mathbf{y}_{1:k}) = \frac{f(\mathbf{y}_{1:k} | \mathbf{h}_{0:k}) f(\mathbf{h}_{0:k})}{\int f(\mathbf{y}_{1:k} | \mathbf{h}_{0:k}) f(\mathbf{h}_{0:k}) \, d\mathbf{h}_{0:k}} \quad (9)$$

The problem amounts to finding a transform $f(\mathbf{h}_{0:k+1} | \mathbf{y}_{0:k+1}) = \Phi\{f(\mathbf{h}_{0:k} | \mathbf{y}_{0:k})\}$. To find out the required transform Φ we invoke Bayes' theorem:

$$\begin{aligned} &f(\mathbf{h}_{0:k+1} | \mathbf{y}_{1:k+1}) = f(\mathbf{h}_{0:k+1} | \mathbf{y}_{k+1}, \mathbf{y}_{1:k}) = \\ &\tfrac{f(\mathbf{y}_{k+1} | \mathbf{h}_{0:k+1}, \mathbf{y}_{1:k}) f(\mathbf{h}_{0:k+1} | \mathbf{y}_{1:k})}{f(\mathbf{y}_{k+1} | \mathbf{y}_{1:k})} \end{aligned} \quad (10)$$

However, the observations are conditionally independent: $f(\mathbf{y}_{k+1} | \mathbf{h}_{0:k+1}, \mathbf{y}_{1:k}) = f(\mathbf{y}_{k+1} | \mathbf{h}_{0:k+1})$. This leads to the required recursive formula:

$$f(\mathbf{h}_{0:k+1} | \mathbf{y}_{1:k+1}) = \frac{f(\mathbf{y}_{k+1} | \mathbf{h}_{0:k+1}) f(\mathbf{h}_{0:k+1} | \mathbf{y}_{1:k})}{f(\mathbf{y}_{k+1} | \mathbf{y}_{1:k})} \quad (11)$$

The above step updates the prior density $f(\mathbf{h}_{0:k+1} | \mathbf{y}_{1:k})$ once the measurements $\mathbf{y}_{k+1}$ become available. To complete the recursions the prior density has to be specified. This is known as a prediction step:

$$f(\mathbf{h}_{0:k+1} | \mathbf{y}_{1:k}) = \int f(\mathbf{h}_{k+1} | \mathbf{h}_k) f(\mathbf{h}_k | \mathbf{y}_{0:k}) \, d\mathbf{h}_k \quad (12)$$

Equations (11) and (12) constitute a basis for Bayesian recursive estimation. Deceptively, the above recursions are straightforward to perform. However, the integrals involved are, in general, too difficult to compute. An exception is the case where the states evolve according to some linear function and both the state and the observation noise are Gaussian.

B. Application to channel estimation in SFTC

It is well known that the Kalman filter is an optimum Bayesian recursive estimator when both the state transitions and observation systems are linear and both the state and the observation noise are Gaussian. The Kalman filter performs the recursions from previous section, when the underlying problem has a form:

The estimated state $\mathbf{h}_k$ evolves according to:

$$\mathbf{h}_{k+1} = \mathbf{A}_{k+1} \mathbf{h}_k + \mathbf{w}_{k+1} \quad (13)$$

And the observed signal is given by:

$$\mathbf{y}_{k+1} = \mathbf{C}_{k+1} \mathbf{h}_{k+1} + \mathbf{v}_{k+1} \quad (14)$$

The state noise $\mathbf{w}_k$ and the observation noise $\mathbf{v}_k$ are distributed according to:

$$f(\mathbf{w}_k) = N(\mathbf{0}, \mathbf{Q}) \quad (15)$$

$$f(\mathbf{v}_k) = N(\mathbf{0}, \mathbf{R}) \quad (16)$$

Equations (13,14,15,16) imply that the estimated process evolves sequentially and constitutes what is known as Gauss-Markov random process.

We can now cast our problem of channel estimation in SFTC onto the above framework. In order to apply the Kalman filter framework we need to redefine equation (6) to the equivalent form as in (14), which is achieved by defining $\mathbf{C}_k$ and $\mathbf{h}_k$ as:

$$\mathbf{C}_k = \begin{bmatrix} \mathbf{c}_k^T & \mathbf{0}^T & \cdots & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{c}_k^T & \cdots & \mathbf{0}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}^T & \mathbf{0}^T & \cdots & \mathbf{c}_k^T \end{bmatrix} \quad (17)$$

$$\mathbf{h}_k = vec\{\mathbf{H}_k^T\} \quad (18)$$

As mentioned in the previous section, sequential estimation amounts to repeated calculations of two alternating steps: prediction and up-date.

1) Prediction: Suppose that the random variable $\mathbf{h}_k$ conditioned on the observations $\mathbf{y}_{1:k}$, is Gaussian:

$$f(\mathbf{h}_k | \mathbf{y}_{1:k}) = N(\mu_k, \mathbf{P}_k) \quad (19)$$

From equation (13) it can be deduced that $f(\mathbf{h}_{k+1} | \mathbf{h}_k) = N(\mathbf{A}\mathbf{h}_k, \mathbf{Q})$ ($\mathbf{A}$ is a constant matrix here: $\mathbf{A}_k = \mathbf{A}$). Then from equation (12) the predictive marginal distribution is given by (with some abuse of notation):

$$f(\mathbf{h}_{k+1} | \mathbf{y}_{1:k}) = \int N(\mathbf{A}\mathbf{h}_k, \mathbf{Q}) N(\mu_k, \mathbf{P}_k) \, d\mathbf{h}_k \quad (20)$$

After some tedious but straightforward manipulations[1] this becomes:

$$f(\mathbf{h}_{k+1} | \mathbf{y}_{1:k}) = N\left(\mathbf{A}\mu_k, \mathbf{Q} + \mathbf{A}\mathbf{P}_k\mathbf{A}^H\right) \quad (21)$$

The following definitions are made: $\mu_{k+1|k} = \mathbf{A}\mu_k$ and $\mathbf{P}_{k+1|k} = \mathbf{Q} + \mathbf{A}\mathbf{P}_k\mathbf{A}$. Then the predictive density is fully defined by:

$$f(\mathbf{h}_{k+1} | \mathbf{y}_{1:k}) = N\left(\mu_{k+1|k}, \mathbf{P}_{k+1|k}\right) \quad (22)$$

2) Up-date: Using (22) and (16), the update formula of (11) can be specified as (again with some abuse of notation):

$$f(\mathbf{h}_{k+1} | \mathbf{y}_{1:k+1}) = \frac{N(\mathbf{C}_{k+1}\mathbf{h}_{k+1}, \mathbf{R}) \, N\left(\mu_{k+1}, \mathbf{P}_{k+1|k}\right)}{f(\mathbf{y}_{k+1} | \mathbf{y}_{1:k})} \quad (23)$$

After some tedious algebraic manipulations the posterior marginal density is expressed as:

$$f(\mathbf{h}_{k+1} | \mathbf{y}_{1:k+1}) = N\left(\mu_{k+1}, \mathbf{P}_{k+1}\right) \quad (24)$$

with the following notation:

$$\begin{aligned} &\mathbf{P}_{k+1} = [\mathbf{I} - \mathbf{K}_{k+1}\mathbf{C}_{k+1}] \, \mathbf{P}_{k+1|k} \\ &\mu_{k+1} = \mu_{k+1|k} + \mathbf{K}_{k+1}\left[\mathbf{y}_{k+1} - \mathbf{C}_{k+1}\mu_{k+1|k}\right] \\ &\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k}\mathbf{C}_{k+1}^H\left[\mathbf{R} + \mathbf{C}_{k+1}\mathbf{P}_{k+1|k}\mathbf{C}_{k+1}^H\right]^{-1} \end{aligned} \quad (25)$$

Since both the predictive density $f(\mathbf{h}_{k+1} | \mathbf{y}_{1:k})$ and the up-dated posterior density $f(\mathbf{h}_{k+1} | \mathbf{y}_{1:k+1})$ are Gaussian, the mean and covariance describes them completely.

[1]The exact algebraic manipulations involve expanding the two Gaussian densities, completing the square and integrating.

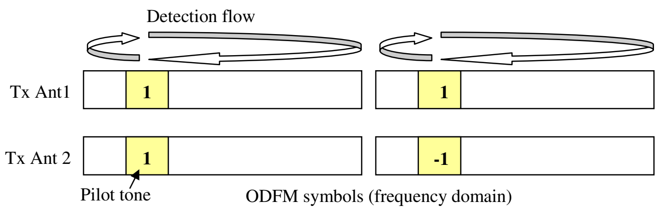

Fig. 2. Detection flow schematic.

C. Description of the proposed algorithm

In the semi-blind case we assume that standard channel estimation is performed on one sub-carrier. This is depicted in figure 2. We assume that the space-frequency code uses two transmit antennas. To avoid any ambiguity at least two encoded OFDM symbols (the SFTC code) are sent from each antennas. The same frequency tone (denoted as $k = 0$) is chosen in each OFDM symbol to carry an orthogonal training sequence. The training results in an initial estimate $\left\{\hat{\mathbf{H}}_0\right\}$. This estimate, together with the corresponding covariance matrix, is propagated to the neighbouring frequency tone using eq (21). The prior channel estimate ($\hat{\mathbf{H}}_{k+1|k}$ and thus $\hat{\mathbf{h}}_{k+1|k}$ via eq. (18)) is simply the mean ($\mu_{k+1|k}$) of the predictive density in (21). It is crucial for the success of this method that the SFTC trellis always starts from a known state. This is depicted in figure 3, where it is assumed that the trellis starts from a zero state. This is a fairly common assumption in trellis coding, typically done to improve the performance (however not essential as it is here). For simplicity, of presentation figure 3 depicts a 4 state bpsk space-frequency code. In this case the two possible input symbols ('0' and '1') would result in two transitions (to state 0 and state 1 respectively). Corresponding to the two transitions there are two codewords: $\mathbf{c}^{(0,0)}$ and $\mathbf{c}^{(0,1)}$ (and thus $\mathbf{C}^{(0,0)}$ and $\mathbf{C}^{(0,1)}$ via eq. (17)). The superscript (i, j) denotes a transition from an i^{th} state to the j^{th}. Subsequently, the update of the channel estimates is performed. Using set of equations (25) the process (channels), the channels covariance matrices and the Kalman gain matrices all get updated. Since the $\mathbf{C}^{(i,j)}$ are in general different, the update process results in obviously different posterior estimates for states 0 and 1.

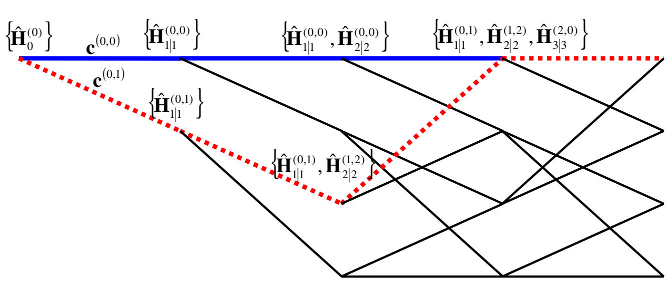

Fig. 3. Trellis description of the algorithm.

This procedure continues until the state transitions in the trellis merge (the fourth segment of the trellis in figure 3) and all hypotheses till then are retained. At this stage a decision is made. The two paths merging at each stage correspond to two distinct hypotheses. Each with a set of codewords $\left\{\mathbf{c}_{1:k}^{(i,j)}\right\}$ and a set of $\left\{\hat{\mathbf{H}}_{1:k}^{(i,j)}\right\}$. Assuming that the Kalman filters track the channels with sufficient accuracy, a decision can be made to retain only one hypothesis using a Euclidean distance criterion eq (7). For example, in figure 3 the dashed path is retained and with it the history of channel posterior estimates $\left\{\hat{\mathbf{H}}_{1|1}^{(0,1)}, \hat{\mathbf{H}}_{2|2}^{(1,2)}, \hat{\mathbf{H}}_{3|3}^{(2,0)}\right\}$. This is the last estimate in this set that will be propagated to obtain the prior estimate for all transitions originating from this state. This procedure is repeated for all states and all frequency tones. If the trellis is terminated (forced to the zero state) the last decision taken at the zero state will identify a path assumed to be correct and with it the whole space-frequency codeword $\left\{\mathbf{c}_{1:K}^{(i,j)}\right\}$ and the channel estimates $\left\{\hat{\mathbf{H}}_{1:K}^{(i,j)}\right\}$.

In the blind case there is no initial training and the initial estimate is set to zero $\left\{\hat{\mathbf{H}}_0 = \mathbf{0}\right\}$.

Table I summarises the proposed algorithm.

TABLE I
ALGORITHM SUMMARY

INITIALISE: $\hat{\mathbf{h}}_0, \mathbf{A}, \mathbf{Q}, \mathbf{P}_0$

RECURSIONS: for *k=1:K*, for *j=1:J*

$$\hat{\mathbf{h}}_{k+1|k}^{(j)} = \mathbf{A}\hat{\mathbf{h}}_{k|k}^{(i)}$$
$$\mathbf{P}_{k+1|k}^{(j)} = \mathbf{Q} + \mathbf{A}\mathbf{P}_{k|k}^{(i)}\mathbf{A}$$
$$\left[\mathbf{C}_{k+1}^{(j)}, \psi_{k+1}^{(j)}\right] = \arg\min_{\tilde{\mathbf{C}}^{(i,j)}} \left\{ \left\| \mathbf{y}_{k+1} - \tilde{\mathbf{C}}^{(i,j)}\hat{\mathbf{h}}_{k+1|k}^{(j)} \right\|^2 + \psi_k^{(i)} \right\}$$
$$\mathbf{K}_{k+1|k+1}^{(j)} = \mathbf{P}_{k+1|k}^{(j)}\mathbf{C}_{k+1}^{(j)\,H}\left[\mathbf{R} + \mathbf{C}_{k+1}^{(j)}\mathbf{P}_{k+1|k}^{(j)}\mathbf{C}_{k+1}^{(j)\,H}\right]^{-1}$$
$$\mathbf{P}_{k+1|k+1}^{(j)} = \left[\mathbf{I} - \mathbf{K}_{k+1|k+1}^{(j)}\mathbf{C}_{k+1}^{(j)}\right]\mathbf{P}_{k+1|k}^{(j)}$$
$$\hat{\mathbf{h}}_{k+1|k+1}^{(j)} = \hat{\mathbf{h}}_{k+1|k}^{(j)} + \mathbf{K}_{k+1|k+1}^{(j)}\left[\mathbf{y}_{k+1} - \mathbf{C}_{k+1}^{(j)}\hat{\mathbf{h}}_{k+1|k}^{(j)}\right]$$

TERMINATE:

trace back $\left\{\mathbf{c}_{K:1}^{(0)}, \hat{\mathbf{h}}_{K:1}^{(0)}\right\}$

IV. SIMULATION RESULTS

The performance of the proposed algorithm is investigated using simulated MIMO-OFDM and the 16 state 4 PSK space-time code of [8] (used here as space-frequency code). A 64 point FFT is used with all subcarriers occupied. A frame is constructed from 126 information symbols that are encoded to a space-frequency code-word. When combined with one pilot tone in each OFDM symbol, the span is two OFDM symbols. The pilot tones are placed at the beginning of each OFDM symbol and each OFDM symbol includes a cyclical prefix of 16 symbols. For simulation purposes we use a simple channel with L = 3 taps, all i.i.d. complex circular Gaussian ($N\left(0, (2L)^{-1}\right)$ per dimension). In frequency domain the channels are modelled as random walk: ($\mathbf{A} = \mathbf{I}$ and $\mathbf{Q} = .05\mathbf{I}$).

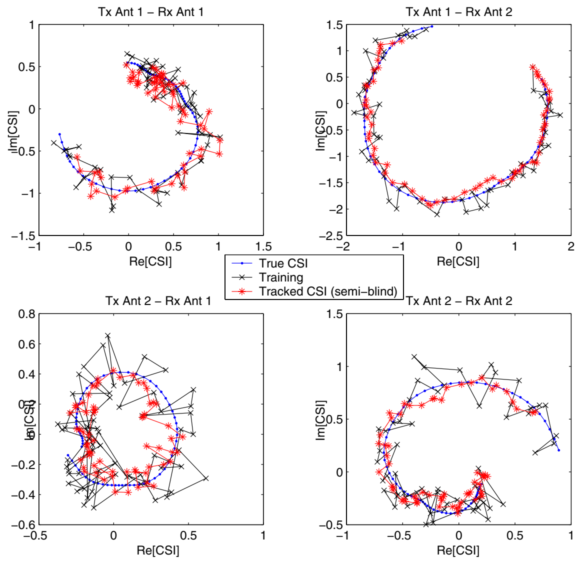

Fig. 4. Channel estimation and tracking by the proposed technique: 64 FFT size, one (first) trained tone, frame length 126 symbols (2 OFDM symbols), 2 Tx by 2 Rx antennas, SNR = 15 dB per receive antenna, 16 state 4-PSK (BBH) code.

Figure 4 depicts a typical set of CSI in the frequency domain together with the estimate provided by the proposed technique corresponding to a correct hypothesis. For comparison, the performance of the proposed techniques is checked against a trained version of the same architecture. In the trained version, prior to the space-frequency code transmission, training sequences are sent. The training consists of a sequential transmission of preambles (1 OFDM symbol per antenna). As can be seen, the proposed technique closely tracks the channel realisation. Quantitative results are presented in figures 5 and 6 where average frame error rate and an ensemble-averaged mean squared error of the channel estimate are depicted respectively. As expected, both versions of the proposed technique loose some diversity gain as compared to the trained version. However, they also deliver reliable SFTC detection and MIMO channel estimation. In this example the bandwidth efficiency of the novel techniques is improved by a factor of 100%, since the frame in the trained technique consists of 4 OFDM symbols (2 data + 2 training).

V. Conclusions

A novel joint detection and channel estimation technique has been developed for trellis coded MIMO-OFDM. The technique uses only a single initial estimate in the semi-blind and none in the blind mode. In both cases the entire channel estimate is fully recovered and SFTC is decoded. Clearly, the major appeal is that there is no need for a training sequence at the expense of increased complexity and slightly worse performance. The technique can also be applied to Space-time trellis coded single carrier systems, where even fast fading channels can be efficiently tracked.

Acknowledgments

The authors wish to thank Dr. Christophe Andrieu for very insightful comments and suggestions, and Toshiba TREL. Ltd. for sponsoring the work presented in this article.

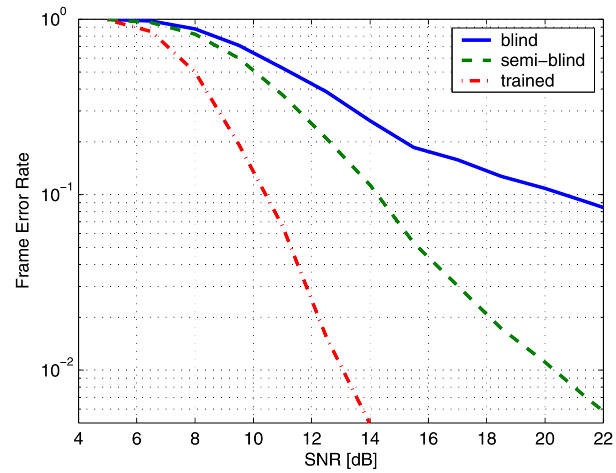

Fig. 5. Frame error rate performance of the proposed techniques, 2 by 2, 16 state 4 PSK code.

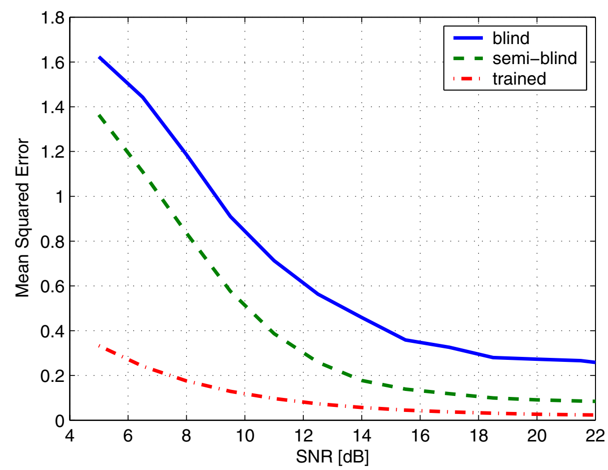

Fig. 6. Ensemble-averaged mean squared error of the channel estimate for the proposed techniques, 2 by 2, 16 state 4 PSK code.

References

[1] G.J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311 – 335, 1998.

[2] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication - performance criterion and code construction," *IEEE Trans on Information Theory*, vol. 44, no. 2, pp. 744 – 765, 1998.

[3] D. Agraval, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded ofdm for high date rate wireless communications over wideband channels," *in Proc. 48th IEEE VTC, Ottawa, Canada*, 1998.

[4] Z. Liu, X. Ma, and G. Giannakis, "Space-time coding and kalman filtering for time selective fading channels," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 183 – 186, 2002.

[5] J. Zhang and P. Djuric, "Joint estimation and decoding of space-time trellis codes," *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 3, pp. 305 – 315, 2002.

[6] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Transactions on Communications*, vol. 42, no. 234, pp. 1000 – 1011, 1994.

[7] R. Raheli, Polydoros, and C Tzou, "Per-survivor processing: A general approach to mlse in uncertain environments," *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 354 – 364, 1995.

[8] S. Baro, G. Bauch, and A. Hansmann, "Improved codes for space-time trellis coded modulation," *IEEE Communications Letters*, vol. 4, no. 1, 2000.